Update

Experiences with Accelerators and Coprocessors in High-Performance Computing: HPC User Forum, September 15–17, 2014, Seattle, Washington

Robert Sorensen
Earl C. Joseph, Ph.D.
Steve Conway
Chirag Dekate, Ph.D.

IN THIS UPDATE

This IDC update captures part of the proceedings at the 54th HPC User Forum held in Seattle, Washington. HPC research facilities around the world are actively exploring the utility of accelerator and coprocessors in their high-performance computing (HPC) systems as well as some of the programming challenges of realizing the maximum performance potential of these components. Paul Muzio, the new chairman of the HPC User Forum, moderated this session, which had both industry and academic presenters summarizing their experiences with accelerators and coprocessors in their HPC environments. The speakers at this accelerator/coprocessors session included Seid Koric, NCSA, speaking about GPUs and industrial achievements; Edoardo Apra, Pacific Northwest National Laboratory (PNNL), speaking on the development of Intel MIC codes in NWChem; Troy Porter, Stanford University, speaking on experiences with the Intel Phi and astrophysics; and John Turner, ORNL, talking about accelerators used for R&D related to light water reactors.

Accelerators and coprocessors offer the potential for breakthrough performance gains in a number of critical academic, government, and commercial applications, but there is still significant work to be done porting existing codes and generating new ones that can fully capture the computational capabilities of these devices. IDC believes that in the near term, continued progress in accelerator/coprocessor hardware capabilities, along with increasingly sophisticated software development tools, will yield additional performance payoffs and attract a growing base of users.

Industrial Achievements on Blue Waters Using CPUs and GPUs — Seid Koric, NCSA

Tom Hughes, 2003: "Computers will always come in one size and one speed: too small and too slow." NCSA industrial partner Rolls Royce has an engine simulation problem needing 1T degrees of freedom. We are probably the largest HPC industrial engagement in the world. The National Petascale Computing Facility: 88,000 sq ft and 25MW. We are using 40-50% of this now. It's now 100Gbps, going to 400Gbps. We have many industrial partners.

- NCSA supercomputers:
  - **iForge:** Dedicated to industrial users and on-demand use, with reservation and hosting options

- **Blue Waters:** 1PF sustained on real applications; 400,000 x86 cores; 12 million CUDA cores; 1.5PB RAM; 400PB+ storage

As you scale, communication becomes dominant, versus computation.

- GPU options:
  - Libraries for drop-in acceleration
  - OpenACC to incrementally accelerate (We started this in 2011. It looks like OpenMP. The name of the game for GPGPUs is to minimize data movement.)

Performance limitations for scaling MPI codes on GPUs: You have to go through the PCI to reach the MPI. Recently, there's been a GPU-aware MPI from NVIDIA and also a Cray MPI. Then there's unified virtual memory where you don't have to worry much about where the data is stored.

MPI plus OpenACC example: A hybrid Laplace 2D equation with FED – MPI plus OpenMP versus GPU-aware MPI plus OpenACC. It's about 3x faster with the latter. ISV multinode GPU acceleration on XK7 on Abaqus also provided a good speedup.

Production engineering work using GPUs at NCSA is close to zero. It's all experimental/exploratory. It will take 5–10 years for GPUs to go into production use for important ISV codes.

Development of Intel MIC Codes in NWChem — Edoardo Apra, Pacific Northwest National Laboratory

EMSL is a national scientific user facility within Pacific Northwest National Laboratory in Richmond, Washington. It is funded by the DOE Office of Science's Office of Biological and Environmental Research. NWChem was developed at EMSL for molecular science. It was designed for MPP systems and today runs on a wide range of computers. It's open source and has a broad user community. It uses global arrays/ARMCI for parallelization.

NWChem uses high-accuracy methods in molecular science, including the study of the energetics. We adapted NWChem for heterogeneous architectures. We are application scientists, not computer scientists. We will be developing more algorithms for heterogeneous systems.

Global arrays are distributed dense arrays that can be accessed through a shared memory-like style. Tensor contraction engine (TCE) is a computer-generated code expressed in DSL language, for coding complicated tensor expressions.

Non-iterative EOMCC code: 94% parallel efficiency using 120,000 cores at ORNL. GPU CCSD benchmarks: Titan Cray XK7 at ORNL – 98 nodes, 8 cores/node plus 1 GPU. It's written in CUDA. We achieved a speedup on GPUs versus CPUs of 5.6x.

Xeon Phi port NWChem for design decisions: The kernel CCSD is suitable for coupled cluster method to study total energy in a molecule. There are many opportunities for data reuse. Minimizing data transfers has been successfully used before for a GPU port. We adopted an offload approach. We first used OpenMP target offloading and then moved to LEO. We did not modify the existing GA

parallelization structure; we only localized changes to the existing FORTRAN source code. Cons: Adoption of a proprietary development tool called LEO. Wish: Will OpenMP/OpenACC/LEO eventually merge into a standard for accelerators?

The experimental setup used 1 of the 16 cores on each node. Four ranks were dedicated to offload, and four to do CPU work (number crunching). Performance: Xeon plus Phi beat either of these alone out to 62,560 threads on CCSD on a small benchmark. It was the same with a larger benchmark.

Future potential: Native rather than offload mode would take more work but might be more rewarding. Substantial work is needed by the developer. The compiler helps in some circumstances but not in others.

Double Rewards of Porting Scientific Applications to the Intel MIC Architecture — Troy Porter, Stanford University

Astrophysics and particle physics are my research areas. The Xeon and Phi architectures and programming models are similar. It's true that Phi is programmed like Xeon. About 75% of Phi cores are usable in practice for our codes. Native programming model for Phi: The app runs on Phi, and the Xeon, if used, acts as a coprocessor. The other approach is offload mode, where the app runs on Xeon and communicates with the coprocessor to offload parts of the work.

Case study: Building a 3D model of the Milky Way galaxy using 2D sky surveys. I'm interested in interstellar dust, which is important for the formation and lifetimes of galaxies. The challenge is to build a 3D model from various 2D views. The code is FRaNKIE, which uses an inferencing technique (Bayesian inference). Challenge: Modeling the stochastic heating of cosmic dust by starlight is very time consuming and would have taken hundreds of years of CPU time to run. It is therefore intractable. We use the MultiNest Bayesian analysis engine with the FRaNKIE radiation transport Monte Carlo code plus HEATCODE for cosmic dust heating. HEATCODE goes to the Xeon Phi. We optimized the problem for this and went from hundreds of CPU years to hundreds of CPU days, meaning a month or so running on our moderate-size cluster, which is tractable. Input: incident electromagnetic radiation field. Output: spectrum of reemitted photons.

Optimization road map for doing this:

- Scalar optimization
- Vectorization: By default, the compiler does not assume alignment. This must be hinted to the compiler, as with automatic additional vectorization). 512-bit vector holds 16 single-precision FP numbers.
- Thread scalability: We are using an OpenMP parallel region inside of the #pragma offload. We modified the library interface. We reduced the per-thread memory footprint because 240 threads do not fit into Xeon Phi memory.
- Memory access
- Communication

We used the Intel suite of tools, VTune. It's really, really nice to use because it allowed us to identify bottlenecks in the code. After optimization, performance was 620x faster than Xeon and also 125x faster on the Xeon. So the effective acceleration factor is 1.9x. One code was used for both optimizations and runs on both platforms. We also got improvements in compute density and power efficiency.

Future-proofing applications for Knights Landing: from 22nm PCIe coprocessor to 14nm standalone processor.

Summary

You get double rewards using the same code and optimization for MIC and CPU. It doesn't require much work. It took about one month by a post doc and two more months to get everything we got out of it. I'm very impressed with Intel MIC.

Much of what we did for our optimization would apply to other codes in other areas. Knights Landing offers the potential for use as a heterogeneous standalone processor.

Experiences Using Accelerators at ORNL — John Turner, ORNL

I really enjoyed the previous talk. One way to get me upset is to tell me Roadrunner was a failure, and the Cell processor was a dumb idea that got cancelled. Modifying codes for Cell also sped them up on x86 processors, as Porter said. That's been our experience with modifying codes for GPUs at Oak Ridge, too.

DOE Leadership Computing Program: two machines – a Blue Gene at ARL and Titan at ORNL. 60% of the access goes to the INCITE program, 30% to ASCR, and 10% is discretionary.

Titan: Cray XK7 with AMD Opteron plus NVIDIA Tesla, 27.1PF peak, 24.5 GPU plus 2.6 CPU

The Center for Accelerated Application Readiness (CAAR) was created as part of the Titan project. We created teams and selected applications, including combustion, astrophysics, climate, and molecular dynamics.

For each application, we had an OLCF person, a Cray engineer, an NVIDIA developer, and others.

- Multiple acceleration methods were explored.
- Speedups Titan heterogeneous system versus non-heterogeneous Cray (Jaguar)
- LAMMPs had a speedup of 7.4x.
- S3D 2.2
- Denovo 3.8 (optimized code also ran 2x faster on the Cray XT5 Jaguar) (This code scales beyond 200,000 cores.)
- WS-LSMS 3.8 had 7.3x less power consumption.

GPU usage has increased within the INCITE program.

CASL is a DOE innovation hub. The full name is Consortium for Advanced Simulation of Light Water Reactors. It's the only innovation hub devoted entirely to simulation. Goal: Bring HPC-based simulation back to the nuclear reactor design community. Westinghouse, TVA, and the Electric Power Research Institute are the industry partners.

Early Test Stand deployment is already producing benefits for CASL and users:

- Better code installation
- Input processing for heterogeneous codes
- Reductions in user problem setup times
- Analysis of new design features

Westinghouse VERA Test Stand: VERA was deployed on a Westinghouse cluster for a high-impact industrial application.

AP1000 advanced first core model: It's a very complicated nuclear reactor core, so a good thing to analyze. The analysis was completed by Westinghouse in January 2014, not using Denovo or Monte Carlo. But since then, three methods were compared for analyzing the cores using Titan: SPn, Sn, and Monte Carlo. In using VERA, we ran some of the largest Monte Carlo calculations ever run (1 trillion particles), with excellent agreement with KENO-IV – the older code.

Westinghouse was very happy, and this work received an IDC HPC Innovation Excellence Award. That was very important to Westinghouse. Acceleration efforts are underway for other VERA components.

LEARN MORE

Related Research

Additional research from IDC in the technical computing hardware program includes the following documents:

- *Worldwide Broader HPC 2014-2018 Forecast: Servers, Storage, Software, Middleware, and Services* (IDC #248835, June 2014)
- *When Massive Data Never Becomes Big Data* (IDC #lcUS24922014, June 2014)
- *Worldwide Technical Computing Server 2014-2018 Forecast* (IDC #248779, May 2014)
- *Perspectives on High-Performance Data Analysis: The Life Sciences* (IDC #248348, May 2014)
- *Global HPC Market Dynamics in 2013* (IDC #248137, April 2014)

- *Industrial Partnership Programs and High-Performance Computing: HPC User Forum, April 7-9, 2014, Santa Fe, New Mexico* (IDC #248113, April 2014)
- *Disruptive Technologies in High-Performance Computing: HPC User Forum, April 7-9, 2014, Santa Fe, New Mexico* (IDC #248112, April 2014)
- *Advances in Processors, Coprocessors, and Accelerators in High-Performance Computing: HPC User Forum, April 7-9, 2014, Santa Fe, New Mexico* (IDC #248111, April 2014)
- *International Perspectives on Industrial High-Performance Computing Partnerships: HPC User Forum, April 7-9, 2014, Santa Fe, New Mexico* (IDC #248122, April 2014)
- *Worldwide HPC Public Cloud Computing 2014-2017 Forecast* (IDC #247846, April 2014)
- *Summary of IDC's 2014 Research in the Use of HPC by Oil and Gas Organizations* (IDC #247704, March 2014)
- *IBM Sale to Lenovo Opens Opportunity for Other HPC Vendors* (IDC #lcUS24694314, February 2014)
- *IDC's Worldwide High-Performance Computing Predictions 2014* (IDC #WC20140211, February 2014)
- *Seagate Looking for the X Factor in Its Acquisition of Xyratex* (IDC #lcUS24555413, December 2013)
- *Micron Demonstrates Technologies to Address Emerging Challenges in Big Data Applications* (IDC #244843, December 2013)
- *Market Analysis Perspective: Worldwide HPC, 2013 – Directions, Trends, and Customer Requirements* (IDC #244742, December 2013)
- *HPDA Pulse: 2013 Software and Consulting Market Analysis* (IDC #244513, November 2013)
- *HPDA Pulse Results: 2013 Hardware and Storage Market Analysis* (IDC #244493, November 2013)
- *HP FY13: Revenue Declines Abate on Stronger Core Business* (IDC #lcUS24466413, November 2013)
- *Catalyst Supercomputer Heralds Shift to More Balanced Architectures* (IDC #lcUS24437513, November 2013)
- *China Eyes 10,000-Fold Data Reduction for Internet of Things* (IDC #lcUS24392513, October 2013)
- *HPC User Forum, October 2013, Seoul, Korea* (IDC #243786, October 2013)
- *Tools and Techniques for Technical Computing in Life Sciences: HPC User Forum, September 2013, Boston, Massachusetts* (IDC #243778, October 2013)
- *Perspectives on Quantum Computing: HPC User Forum, September 2013, Boston, Massachusetts* (IDC #243777, October 2013)
- *National and International Initiatives: HPC User Forum, September 2013, Boston, Massachusetts* (IDC #243776, October 2013)
- *Issues in High-Performance Computing: HPC User Forum, September 2013, Boston, Massachusetts* (IDC #243775, October 2013)

- *High-Performance Data Analysis in the Life Sciences: HPC User Forum, September 2013, Boston, Massachusetts* (IDC #243774, October 2013)
- *Chinese Research in Processor Designs for High-Performance Computing and Other Uses* (IDC #243502, October 2013)
- *World's Fastest Supercomputer Set to Reach Customer in October 2013* (IDC #IcUS24300913, September 2013)
- *The Broader HPC Market 2012-2017 Forecast: Servers, Storage, Software, Middleware, and Services* (IDC #242742, August 2013)
- *IDC's Worldwide Technical Server Taxonomy, 2013* (IDC #242725, August 2013)
- *China Regains Top Supercomputer Title* (IDC #IcUS24190613, June 2013)
- *10 Things CIOs Should Know About High-Performance Computing* (IDC #241565, June 2013)
- *Worldwide High-Performance Data Analysis 2013–2017 Forecast* (IDC #241315, June 2013)
- *Top Issues for HPC Sites: HPC User Forum, April 29–May 1, 2013, Tucson, Arizona* (IDC #241463, June 2013)

About IDC

International Data Corporation (IDC) is the premier global provider of market intelligence, advisory services, and events for the information technology, telecommunications and consumer technology markets. IDC helps IT professionals, business executives, and the investment community make fact-based decisions on technology purchases and business strategy. More than 1,100 IDC analysts provide global, regional, and local expertise on technology and industry opportunities and trends in over 110 countries worldwide. For 50 years, IDC has provided strategic insights to help our clients achieve their key business objectives. IDC is a subsidiary of IDG, the world's leading technology media, research, and events company.

Global Headquarters

5 Speen Street
Framingham, MA 01701
USA
508.872.8200
Twitter: @IDC
idc-insights-community.com
www.idc.com